DAITC

Global Scientific and Technical Information Network

DAITC/TR-88/007

Gladys A. Cotter
(Defense Applied Information Technology Center)

December 1988

Defense Applied Information Technology Center

1800 North Beauregard Street
Alexandria, Virginia 22311
(703) 998-4787
Fax No. (703) 931-3968

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) DAITC/TR-88/~~005~~ 007	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Defense Applied Information Technology Center 6b. OFFICE SYMBOL (*If applicable*) DTIC-DA	7a. NAME OF MONITORING ORGANIZATION Defense Technical Information Center
6c. ADDRESS (*City, State, and ZIP Code*) 1800 N. Beauregard Street Alexandria, VA 22311-1784	7b. ADDRESS (*City, State, and ZIP Code*) Cameron Station Alexandria, VA 22304-6145
8a. NAME OF FUNDING / SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (*Include Security Classification*)
Global Scientific and Technical Information Network

12. PERSONAL AUTHOR(S)
Gladys A. Cotter

13a. TYPE OF REPORT	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
		8812	8

16. SUPPLEMENTARY NOTATION
Presented at Online 88 Information, 12th International Online Information Meeting, London, England, 6-8 Dec 88

17. COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	Networks, Information Technology, Geatways, International, Interfaces

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

This paper describes the foundations for a global scientific and technical information network which is currently under development at the Defense Applied Information Technology Center (DAITC). The paper highlights advances in information technology which are creating an environment where networking and interoperability strategies are possible on a global basis. Applied information technology modules which have been created to support the Global Scientific and Technical Information Network are described. The network is purposefully designed to incorporate both end-users and their information management intermediaries in a complementary manner making them resolute partners in the work and its rewards.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Gladys A. Cotter	22b. TELEPHONE (*Include Area Code*) (703) 998-4693 22c. OFFICE SYMBOL DTIC-DA

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

GLOBAL SCIENTIFIC AND TECHNICAL INFORMATION NETWORK

G.A. Cotter, Defense Applied Information Technology Center, USA

Keywords: *Networks, Information Technology, Gateways, International, Interfaces*

Abstract: This paper describes the foundations for a global scientific and technical information network which is currently under development at the Defense Applied Information Technology Center (DAITC). The paper highlights advances in information technology which are creating an environment where networking and interoperability strategies are possible on a global basis. Applied information technology modules which have been created to support the Global Scientific and Technical Information Network are described. The network is purposefully designed to incorporate both end-users and their information management intermediaries in a complementary manner making them resolute partners in the work and its rewards.

1. NETWORK OVERVIEW

Scientific and Technical Information (STI) is being produced, enhanced and stored around the globe. Single countries in some cases are acknowledged leaders in select scientific and technical disciplines. But in today's interdisciplinary and interrelated environment, it is almost impossible to cordon off STI needs into concise, neatly bounded subject specialties. Most scientist and engineers need access to STI on a global basis in order to maintain a state-of-the-art awareness. Global economies dictate that every effort be made to reduce unnecessary product and service development costs. This makes it necessary to ferret out and exploit to its full potential globally scattered STI which can reduce these costs.

The United States (U.S.) Defense Technical Information Center (DTIC) is laying the foundations for a global scientific and technical information network (STINET) to service its user community. The purpose of the network is to facilitate Department of Defense (DoD) access to scientific and technical information relevant to DoD mission areas. Key elements of the STINET are DoD managers and scientists; DoD Libraries and Information Analysis Centers (IACs); DoD Laboratories; DTIC; and other DoD, federal, commercial and international databases, systems and sources of STI. DTIC is tasked to provide a focus for molding the STINET, to establish interoperability among the various network components, and to coordinate network evolution. DTIC is to achieve this vision by promoting resource sharing and cooperative efforts and through investigation, experimentation, and application of advanced information science and technology.

Development of a STINET is a huge and challenging task which must be accomplished within existing financial and personnel constraints. In order to make visible progress towards the STINET, DTIC had to carefully identify network requirements and evaluate the currently available technology in which applied research could be invested in both the short term and the long term to meet these requirements. Next, DTIC had to settle on an agenda of purposeful steps that would close in on the ultimate goal. Successful development and implementation of the network depends on devising a "doable" plan of work with room for deviation when opportunities for technological acceleration become apparent.

The following elements are required for a successful network:

- o interoperability and interconnection among geographically disperse systems
- o tools such as pointers and menus, to help locate information on a global basis
- o standardized command functions
- o compatible, multifunctional, flexible software for installation at network nodes
- o interconnection with diverse sources of information including government, commercial and international
- o interconnection with diverse types of information including numeric and factual
- o selectivity and data analysis routines
- o improved delivery systems
- o integration of databases and people bases

In addition, DTIC must ensure that the network design incorporates both information end-users (scientists, engineers and their information specialists, etc.) in a complimentary manner. In order for the network to evolve, these two groups need to become resolute partners in the exploration of how new technology can be applied in an actual information service environment.

2. NETWORK DEVELOPMENT

Prototype Approach

The technology needed to meet the requirements of the STINET was at various stages of readiness. Rather than attempting to build an operational system immediately, DTIC decided to develop prototype systems. The prototype approach provided a means for rapidly integrating new technology components into the network for test and evaluation by actual users with functional requirements. It allows a low level of financial-investment and an opportunity to work with test user groups to add, revise or delete capabilities before introducing them into a production environment with operational users. DTIC selected several key and promising areas in which to invest prototype development efforts. The goal of these efforts was to provide the basic foundation for, and functional capabilities of, the STINET.

Having settled on the prototype development model, DTIC moved to establish a suitable environment for this activity. In partnership with several other DoD organizations who had similar needs, the Defense Applied Information Technology Center (DAITC) was established. The DAITC's mission is to facilitate the prototype development, application and introduction of new information technology. The DAITC is organized into technology laboratories which concentrate in the following areas:

Interoperability/Networking
High Density Information Systems
Artificial Intelligence
Video/Voice Systems
Hypermedia
Database/Text Search

Each laboratory is equipped with hardware and software which is available for use in prototype development, test and evaluation. The DAITC also has a powerful computer and communications network which allows interconnection with international communication networks and resources.

DTIC's prototype efforts are primarily being accomplished at the DAITC. The DoD Gateway Information System, the Integrated Bibliographic Information System and the SearchMAESTRO system (described below) provide examples of prototype systems which are being developed at the DAITC as the foundation for STINET.

The DoD Gateway Information System (DGIS)

The DoD Gateway Information system (DGIS) is an intelligent gateway system which provides information access and analysis. The capabilities goal was to link people, information

services, and computers pertinent to the STINET. The technology embodied in the DGIS will provide the key menus, pointers, interoperability, and interconnection within the STINET. To accomplish this, DGIS will have to function as an electronic switch, a communications interface, and a transaction controller. DGIS will require a variety of alternatives, tailored to different user types and needs, for obtaining and distributing information.

DGIS focuses on streamlining the information retrieval and analysis process. This is accomplished by placing the user at the center of a vast information universe consisting of people bases and databases and providing the user with the navigational tools required to pinpoint and arrive at his destination. In terms of databases, the DGIS is designed to provide users with answers to the questions.

WHAT RELEVANT DATABASES EXIST?
HOW DO I ACCESS THEM?
HOW DO I RETRIEVE INFORMATION FROM THEM?
HOW DO I MANIPULATE THE RETRIEVED INFORMATION?

The DGIS provides a single, easy-to-use interface for identifying, accessing, interrogating, and post-processing information from numerous databases relevant to DoD information needs.

In terms of people bases, the DGIS is designed to answer the questions:

WHAT EXPERTISE IS AVAILABLE ON THE NETWORK?
HOW DO I COMMUNICATE WITH EXPERTS?
HOW DO I SHARE INFORMATION WITH COLLEAGUES?

The DGIS acts as an integrated information system which allows human experts, information users, and information resources to exist and interact in harmony.

Development of DGIS is a multi-year, multi-task project. A prototype system has been developed which was designed for a user community including both intermediaries and end-users. Databases accessed are federal, commercial and international. In addition to large, well-known databases and systems, many small, specialized DoD databases will eventually be part of the DGIS.

The basic components of the system are:

A DIRECTORY OF RESOURCES, SUBJECT SEARCHABLE
A COMMON METHOD FOR ACCESSING AND SEARCHING DIVERSE DATABASES
TOOLS FOR DOWNLOADING AND POST-PROCESSING DATA
TOOLS FOR COMMUNICATING WITH A NETWORK OF EXPERTS AND COLLEAGUES

These are described below.

Directory of Resources

The Directory of Resources will include subdirectories with references to databases, people, and computing resources. In the first phase, the Directory of Databases is being developed using the INGRES relational database management system. The Directory of Databases contains information on the content, scope, and availability of selected databases. The Directory is subject-searchable; upon entering a topic of interest, the user is provided with a list of appropriate databases.

The Directory of People will contain references to experts in subject areas and information retrieval techniques who may be contacted via the network. The Directory of Computing Resources will contain references to computing resources, such as supercomputers which can be used for data analysis and modeling, available through the network.

Interfaces for Searching Diverse Databases

One of the primary goals of DGIS is to relieve the user of the need to learn and master separate commands and protocols for each database. As mentioned earlier, the DGIS target user community includes both end-users and intermediaries. DTIC found that end-user and intermediary interface needs are very different. (An expansive natural language interface requiring artificial intelligence applications appealed to both populations, but could not be accomplished with existing technology in the short term.)

A dual approach was adopted for the interface design, incorporating separate strategies for intermediaries and end-users. For the intermediary, a common command language is being developed. To satisfy the end-user, the EasyNet database searching service has been integrated into DGIS under the name SearchMAESTRO. SearchMAESTRO is a menu-driven database front-end which provides access to over 900 commercial databases. Search-MAESTRO access is now an option within DGIS and is described in greater detail in a later section of this paper.

Post-Processing

Information retrieved from databases often requires analysis or post-processing in order to become useful to the researcher. A library of post-processing routines for numeric and bibliographic are available in DGIS. In order to post-process data, the user downloads it into a file on DGIS, translates the data into a common format, and calls up one of several available post-processing routines.

People Bases and Databases

The goal in obtaining information is to acquire knowledge. Much of the information we need resides in the minds of human experts. Therefore, the DGIS has been designed to allow interaction among people, hence, the concept of people bases as well as databases.

Through the DGIS, users will be able to identify and communicate with experts and colleagues and to connect to information resources. As a first step, we have focused on providing the technology to allow such interaction.

Accessing a database/system is accomplished using the CONNECT command. This command provides users with automatic access to information resources. Users do not have to know telephone numbers, Defense Data Network (DDN) locations, passwords, access protocols or logout protocols. The user enters the CONNECT command and a data resources name. DGIS then establishes a connection to the resource and logs the user in. DGIS uses TYMNET, TELENET, DDN, and commercial telephone lines to establish connections.

The CONNECT command can be used to access information centers worldwide. To be eligible to use the CONNECT command to access a resource, a DGIS user establishes a deposit account with DTIC so they can pay for usage, and the DGIS System Administrator grants them connect permission.

Several mechanisms are available for interconnecting people. An electronic mail service is available twenty-four hours a day. Standard electronic mail features such as send, receive, answer and forward are incorporated. Mail messages can be sent simultaneously to multiple addresses and to every member of pre-established mail groups with lengthy documents attached if needed.

In addition, electronic mail can be used to send information downloaded from a database and placed in a file. A user who does not want to do his own database searching can send a search request to an information specialist without leaving his desk. The information specialist can perform the search, download the results to a file, and send the file to the user. Since the data is stored in a file, both the information specialist and the end-user can use post-processing routines for manipulating and analyzing the data.

The LINK command allows two or more users at different locations to link their terminals so that they view the same data display. All users have control over the display and can issue commands. Through the LINK command information specialists and end-users can together perform interactive database searches. The end-user benefits from the specialist's expertise while the specialist benefits from the end-user's immediate feedback.

Prototype and Beyond

A prototype DGIS has been developed and is currently undergoing test and evaluation within the DoD community. There are currently 150 users testing the system. The DGIS prototype is running on a VAX 11/780 using the UNIX operating system, the INGRES database management system, and a PROLOG interpreter package. The DGIS software is being ported to a Pyramid 98X, an ELXSI 6800, a Gould 6050, and Sun Workstations for

benchmarking and performance evaluation. Based on the results of the performance evaluation, a hardware configuration for a production system will be acquired. The production configuration may consist of several machines networked together. For example, the common command language and post-processing routines could be isolated on a back-end machine. DTIC plans to stabilize a version of the DGIS and offer it as a standard DTIC service in October 1988. Prototype development will continue on a separate development machine, and enhanced versions of the DGIS will be made operational at selected intervals.

The DGIS was developed in prototype as an unclassified, minicomputer-based centralized gateway system. As we move this version into operation, we plan to begin development of a distributed, clustered gateway network. Gateway nodes in the network would be made up of centralized, multi-user configurations and intelligent workstations distributed to users. A selection of gateway capabilities would be available on personal workstations. It makes sense, for example, to have automated connection routines, common command languages, and post-processing routines for frequently-used databases available on a personal workstation. The centralized, mini-based node would be utilized to identify and search infrequently-used resources and for post-processing volumes of data which are beyond the capacity of the personal workstation. Clustered gateway nodes consisting of a centralized gateway processor and personal workstations could be based on geographic, organizational, or subject boundaries. The clustered gateway nodes would interconnect and route users to appropriate nodes when necessary. Development of a classified gateway system is also underway.

Integrated Bibliographic Information System (IBIS)

Development and implementation of the DGIS allows DTIC users to connect to, search and analyze data retrieved from diverse unclassified database services in the federal, commercial and international sectors. Development, as planned, of a classified version of the DGIS will make its reach almost limitless. But what the DGIS has not provided is tools for the development of local databases or catalogs of holdings and tools for local collection management.

The DGIS is designed as an intelligent switching mechanism. The resources targeted by the DGIS are already online. The Directory of Resources is the only database central to and created and maintained on the DGIS. This is a basic design philosophy, not to be altered for fear of deflecting the DGIS from its primary focus -- that of being a gateway.

But there was a need to provide a vehicle to automate and manage local information collections which are manually maintained, very valuable, and very difficult for non-local personnel to use. This need is acutely felt by the DoD library community, a key component of STINET.

Therefore, DTIC initiated development of a library automation system responsive to the networking and local collection management needs of DoD libraries. The system would support centralized resource sharing while allowing local processing flexibility. The objective was to permit DoD libraries to make maximum use of existing information, organize this information to meet local needs, and selectively share newly-generated information with other members of the community. The system designed to accomplish this would have to integrate local control for local collection management functions (reference, cataloging, and circulation) with access to the external resources required for reference, shared cataloging, and other network requirements.

The Defense Nuclear Agency (DNA) was selected as the prototype site for the IBIS. The DNA library had a collection of more than 100,000 holdings and required a system supporting multiple users. The IBIS prototype was implemented at DNA on a VAX 11/750 minicomputer.

Scaling Down. The hardware configuration required to support the prototype IBIS at DNA, a minicomputer-based system, was not economical for smaller DoD technical libraries with collection sizes ranging from 5,000 to 75,000 items. A microcomputer-based system was more appropriate for the lower transaction volumes and smaller operating budgets associated with these libraries. Therefore, DTIC initiated an effort to identify and isolate the special

requirements of these smaller DoD libraries and accommodate them. A software survey was performed to identify packages which were suitable for servicing these smaller libraries and, at the same time, were compatible with the gateway software. The U.S. Army Training and Doctrine Command (TRADOC) library was selected as the prototype site for the microcomputer-based version of the IBIS.

Future Plans. The results and experience gained during the test and evaluation of the DNA and TRADOC prototypes have been used to develop the specifications for a competitive procurement of a production system from a commercial source. The production system will be available for purchase and installation by any library on the STINET in 1989. As a result of the dual approach -- small and large libraries -- an IBIS product line which can meet the needs of any DoD library, regardless of collection size and transaction volume will be available.

The IBIS is the first of DTIC's Local Automation Models and will make network-compatible software available for local installation and use. The IBIS is tailored for bibliographic information, and its target community is DoD libraries. Later, Local Automation Model product lines can be tailored for other user communities and different types of data. For example, models tailored for numeric data may be a requirement for some communities. Depending on the software selected for the production IBIS and the non-bibliographic user requirements identified, it may be possible to modify the IBIS software to meet additional needs. If not, DTIC can develop additional network-compatible local automation models specifically for an identified defense requirement.

Successful deployment of the DGIS and IBIS will provide DTIC with a powerful product/service line for the STINET. Through STINET, users would have a mechanism for interconnecting with globally available information systems, computing resources and people. They would also have a network-compatible vehicle for automating and managing local information, selectively sharing that information with other members of the network, and analyzing information from local and remote resources.

SearchMAESTRO

Both DGIS and IBIS were designed to deliver the power and utility required by a broad section of the targeted STINET users. But part of this user community did not require the full power of the DGIS and the IBIS. This group consisted of end-users who wanted to do some of their own information gathering. The users in this market segment had the requirement to scan literature and locate relevant items in their area of interest. They needed an interface that would provide easy access to a variety of databases. In some cases, the users would not be familiar with existing databases, so they required a system that would select a database for them and guide them through the search process. The information needs of these users tended towards fact retrieval -- such as the latest production statistics for a manufacturing company -- or scanning for relevant items -- such as what newspaper articles have been written on a particular subject. Their need was for some information relevant to the subject rather than an exhaustive search of the subject. We refer to them as casual end-users.

For this segment of the community, DTIC introduced SearchMAESTRO. As mentioned earlier, SearchMAESTRO provides access to over 900 databases. SearchMAESTRO can be used directly or through the DGIS. Users who access SearchMAESTRO directly can access databases via a simple-to-use interface. This interface eliminates the need to learn unique database command languages and search techniques. The user has two modes of operation from which to choose. In the first mode, SearchMAESTRO leads the user through a series of questions and answers and select the database for him.

In the second mode, the user can select the database he wants and use the SearchMAESTRO interface for executing the search. With either mode, the user can view search results on the screen, print, or save the results using local equipment. The main reason for accessing SearchMAESTRO through the DGIS is to take advantage of the post-processing and electronic mail utilities. The users must always observe fair use practices when dealing with copyrighted material.

The unclassified portions of the Technical Reports and the Work Unit Information system files of DTIC's Defense Research Development, Test and Evaluation Online system (DROLS) are being made available to registered users through SearchMAESTRO.

Any time man and computer meet, a diversity of problems arise which can best be handled through human intervention. Therefore, SearchMAESTRO provides online user assistance through a function called "SOS" for "Save Our Search". At any point during a SearchMAESTRO session, the user can simply enter "SOS" and a search expert will respond. The search experts are trained to interpret reference questions, be knowledgeable about available sources of online information, and know how SearchMAESTRO works. DTIC supplies the search experts for the DROLS files, and Telebase Systems, Inc. supplies experts for all other systems.

SearchMAESTRO is currently available to all DTIC government users. These users access SearchMAESTRO both directly and through the DGIS.

SearchMAESTRO is just the first of several end-user interface options we plan to offer STINET users. DTIC has sponsored two conferences in the area of interface technology and maintains a constant alert for new offerings. We are especially interested in expert systems and personal computer-based interfaces.

3. NETWORK FUTURE

Baseline STINET services will be available with the implementation of DGIS, IBIS and SearchMAESTRO. Operational users will be offered access to the network in 1989. With operational status network maintenance will become a major factor. The maintenance required for STINET goes far beyond maintenance requirements associated with hardware and software. It is the intellectual maintenance which will be the challenge.

This challenge will involve tracking changes instituted at remote resources so that appropriate modifications can be initiated on the STINET and network harmony will not be disrupted. STINET maintenance will also involve adding new resources to the network. Information specialists will have to determine what resources should be added to the network and in what priority order. Most importantly, information specialists will be called upon to attest to the validity and reliability of data sources. They will have to ensure that external resources they use, or recommend to their users, are credible. They will have to develop policies concerning how locally generated information will be controlled and validated. This will be critical; the use of incorrect information for decision making and planning could prove disastrous to an individual or to an institution.

As the number of resources available on the network grows, users will want tailored "views" of information personalized to meet their needs. Information "boutiques", organized by subject and interconnected, will form a virtual worldwide, multimedia library. Information specialists will be called upon to organize and oversee these global libraries.

New services must be continually added to the STINET baseline so that the network will keep pace with user needs to the extent that economics and technology allow. One of the most useful services which could be added to this global network would be an automatic translation facility. This would be useful in cases where pertinent information is stored in a language which is not native to the user. Automatic translation would make the information rapidly available in a useful format. The automatic translation service could also be used in conjunction with the "people bases." The translation capability could be used to eliminate the language barrier among users on the network. For example, users could send and receive electronic mail messages in the language of their choice. The technology for this type of service is still in its initial development stages. A limited prototype system is currently in the planning stages and will be evaluated by test users next year.

Optical and video technology will also have a place in the network. One of the questions which must be addressed is when to distribute resources on optical media for local use rather than making the resources available through telecommunications. Distributing information on optical media often has economic advantages when secure communications are required. Optical and video technology brings us into the era of "hypermedia" where we have to redefine our conventional definitions of information systems and the data which they contain.

The challenge of building a global scientific and technical information network will stretch far into the future. Network services will be added and deleted based on changes in technology and user requirements. The unchanging factor in successful development of the

network will be the partnership among end-users, information specialists and network developers. As long as this cooperative relationship continues, the STINET will continue to grow.

Gladys A. Cotter
Defense Applied Information Technology Center
1800 N. Beauregard Street
Alexandria, VA 22311-1784 U.S.A.
(703) 998-4600